

正则化方法同化多普勒天气雷达资料及 对降雨预报的影响*

赵延来^{1)†} 黄思训^{1)‡} 杜华栋¹⁾ 仲跻芹²⁾

1) (中国人民解放军理工大学气象学院, 南京 211101)

2) (中国气象局北京城市气象研究所, 北京 100089)

(2010年5月21日收到; 2010年9月26日收到修改稿)

基于三维变分同化(3DVAR)的思想,提出一种适用于多普勒天气雷达资料直接同化的正则化方法.从寻求 $Y^o = H(X)$ 带有偏差 δ 的极小模解出发,引入正则化思想,并给出与3DVAR方案不同的新代价函数.针对2008年8月14日发生在北京地区的一次局地暴雨过程,设计了一组数值试验,并采用L曲线准则后验选取最优正则化参数.数值结果表明:正则化方法和3DVAR方案均能有效同化多普勒雷达资料,雷达径向速度的同化效果明显好于反射率因子;正则化参数取值越小,分析场增量越明显,同化进入初始场的中小尺度信息越多,分析场与观测场越接近;同化得到的分析场,能够在一定程度上缩短模式起转调整时间,提高短期降雨预报的准确率,其中最优选取正则化参数的同化试验能够延长雷达资料同化时效,进一步改善降雨预报效果,合理雨区雨带分布,相应得到的TS评分也更高.

关键词: 多普勒雷达资料, 正则化, 3DVAR, 降雨预报

PACS: 92.60.Wc, 92.60.Qx

1. 引言

强对流性天气如局地暴雨、冰雹、龙卷和大风等往往会造成严重的自然灾害,因其局地性强、时间短、尺度小,物理过程高度非线性,给现有的中小尺度数值预报提出很大的挑战.要提高中小尺度天气的预报水平和强对流性天气预测预警能力,有两个关键的问题要解决:其一是数值模式必须达到处理对流过程的分辨率.随着计算机能力的不断增强,中小尺度数值天气预报模式已经达到了几公里的分辨率.其二是用来初始化模式的观测资料,要能详细描述中小尺度大气的真实状态.对于常规观测资料而言,由于其分辨率低,只能提供大尺度天气系统的信息,难以对中小尺度天气系统进行有效的分析和预报.而多普勒天气雷达观测能够详细刻画中小尺度天气系统内部的风场结构和水汽场信息,是提高临近预报和短时定量降雨预报(QPF)准确率的有效手段.由于多普勒天气雷达观测变量为

径向速度和反射率因子,而非数值模式基本变量.因此如何将雷达观测变量和模式变量联系起来,一直是多普勒雷达资料同化研究的重点.现有的多普勒天气雷达资料同化方法主要可以分为两类:一是间接同化,即基于物理规律和动力学关系,先将雷达观测反演成模式变量后作为常规观测或超级观测进行同化;二是直接同化,即基于现代资料同化方法和特定的观测算子,将雷达非模式变量资料直接输入模式,改善模式初始场.

国内外许多学者一直致力于多普勒天气雷达资料的同化研究. Xiao等^[1]以MM5-3DVAR为平台,用Richardson方程将垂直速度增量包含在三维变分同化(3DVAR)中,有效改善了径向速度的同化效果.随后, Xiao等^[2]又将暖雨参数化过程加入WRF模式的3DVAR系统,直接同化雷达反射率因子,改进了台风路径预报和QPF. Sun和Crook^[3-5]在云模式的基础上建立了多普勒雷达四维变分分析系统(VDRAS),采用连续的四维变分同化(4DVAR)循环过程直接同化多普勒雷达PPI资

* 国家科技支撑计划(批准号:2008BAC37B03)资助的课题.

† E-mail: zhaoyanlaj@gmail.com

‡ 通讯联系人. E-mail: huangxp@yahoo.com.cn

料,实现了对对流风暴的反演和模拟. Tong 和 Xue^[6]采用集合卡尔曼滤波(EnKF)方法对单多普勒天气雷达的径向速度和反射率数据进行了同化试验,由于模式中包含多相态冰微物理过程,其结果能够模拟出暴风雨完整的三维结构. 国内,万齐林等^[7]采用 GRAPES-3DVAR 系统,在云体雷达回波强度守恒的前提下,建立了回波强度和径向速度与水平风场分析变量的观测算子. 杨毅等^[8]提出 3DVAR 结合物理初始化(PI)方法同化径向速度和反射率因子,有效缩短了模式起转时间和改善了短时降雨预报. 盛春岩等^[9]利用 ARPS 模式及其 ADAS 系统直接同化雷达反射率因子和径向速度,研究了多普勒雷达资料对中尺度数值模式短时定量降雨预报的影响. 许小永等^[10]利用 EnKF 方法对模拟的多普勒雷达资料进行了实验. 在对一次梅雨锋暴雨过程的模拟中,EnKF 方法从多普勒天气雷达资料中反演出了暴雨中尺度系统的动力场、热力场和微物理场.

总的来说,4DVAR 和 EnKF 是目前多普勒雷达资料直接同化较为理想的方法,而且两者均在理想和实际个例中都取得了较好的结果,但是其计算代价非常巨大,业务上广泛采用的还是 3DVAR 方案. 本文在传统 3DVAR 方案基础上,提出用反问题中的正则化方法^[11-14]直接同化多普勒天气雷达资料,克服了传统 3DVAR 方案中存在诸多假定的缺点. 我们首先从反问题的角度来分析大气资料同化,在理论上导出了正则化方法的代价函数形式. 针对一次中尺度暴雨个例,设计了一组数值试验,同时采用 L 曲线方法^[15]确定最优正则化参数. 数值结果表明,3DVAR 方案和正则化方法均能有效同化多普勒天气雷达资料,改善模式起转延迟问题,提高了短期降雨预报的准确率;同时最优正则化参数对应的正则化方法相比于 3DVAR 方案同化效果更好,其降雨预报结果与实况更加接近,得到的 TS 评分也比较高.

2. 正则化方法实施

2.1. 正则化方法

3DVAR 的思想是将模式变量和观测变量作为一个有机整体来考虑,通过最小化表征分析场和观测场及背景场之间误差的代价函数,得到一个物理

上和动力上与模式协调一致的最优初始场. 作如下三个最基本的假设:

- 1) 观测算子是线性且估计是无偏的;
- 2) 背景场和观测误差协方差矩阵是正定对称的;
- 3) 背景场和观测是互不相关的.

采用线性无偏最小方差估计^[16],即可得到标准的 3DVAR 代价函数

$$J(X) = J_b + J_o = \frac{1}{2}(X - X_b)^T B^{-1}(X - X_b) + \frac{1}{2}[H(X) - Y^o]^T \times O^{-1}[H(X) - Y^o], \quad (1)$$

其中, X 表示模式分析变量, X_b 表示背景场, Y^o 表示观测, B 和 O 分别表示背景场和观测场误差协方差矩阵,而 H 为联系模式变量和观测变量的观测算子.

根据最近资料分析^[17],现代变分同化系统所得分析场中只有 15% 的信息是由观测资料贡献的,其余 85% 的信息来自于背景场. 导致以上问题的原因主要有:其一实际大气各变量之间并非简单的线性关系,观测算子应该非线性的;其二观测场和背景场之间是否相关以及相关程度如何无法考证. 所以,3DVAR 代价函数中 J_o 和 J_b 的权重均取 1 就显得不再合理,所得分析场 X 也不一定会使泛函取到最小值. 于是我们引入反问题中的正则化方法^[18-21],以期达到抑制背景场,突出观测场的作用.

下面我们从反问题的角度来研究上述问题. 在不考虑 X_b 与 Y^o 是否相关及估计是否无偏的条件下,(1)式就失去了其理论依据. 对于此类问题,我们应该从以下的观点来进行处理,即在观测误差范围内寻找与 X_b 距离最小的解,

$$\|X^\delta - X_b\| = \min \{ \|X - X_b\| \mid \|Y^o - H(X)\| \leq \delta \}, \quad (2)$$

其中, δ 表示观测误差. 记 $\langle \cdot, \cdot \rangle$ 和 (\cdot, \cdot) 分别表示为观测和模式变量所属于的欧氏空间的内积,式中范数分别定义如下:

$$\|Y^o - H(X)\|^2 = \langle Y^o - H(X), O^{-1}[Y^o - H(X)] \rangle, \\ \|X - X_b\|^2 = (X - X_b, B^{-1}(X - X_b)). \quad (3)$$

上述问题的解 X^δ 是存在且唯一的^[22]. 在数学上, X^δ 称之为 $Y^o = H(X)$ 带有偏差 δ 的极小模解. 容易推导出

$$X_\alpha^\delta = X_b + (\alpha B^{-1} + H^T O^{-1} H)^{-1} H^T \times O^{-1}[Y^o - H(X_b)], \quad (4)$$

此问题亦等价于如下代价函数取极小:

$$J^\alpha[X] = J_0 + \alpha J_b = \min! \quad (5)$$

其中, J^α 称为 Tikhonov 泛函, J_b 称为 Tikhonov 稳定函数, α 称为正则化参数. 反问题的理论告诉我们, 解反问题是否成功, 正则化参数 α 的选取至关重要. 记

$$G(\alpha) = \|Y^\alpha - H(X_\alpha^\delta)\|^2 - \delta^2 = 0, \quad (6)$$

可以证明, (6) 式存在唯一的解 $\alpha = \alpha(\delta)$ [23], 即为最优正则化参数, 一般而言 $\alpha(\delta)$ 取值比较小.

正则化方法与 3DVAR 方案根本不同之处在于, J_b 在正则化方法中起着最优调整观测和背景场的权重且对求解过程起稳定作用. 正则化方法适用范围比 3DVAR 方案更广: 其一, 当正则化参数取 $\alpha = 1$ 时, 即对应 3DVAR 方案; 其二, 采用正则化方法可以保证代价函数(5)存在唯一解, 并使得迭代速度加快, 迅速得到全局最优解; 其三, 正则化方法还可以消除分析解的不稳定性; 其四, 正则化方法克服了文献[17]提出的疑团, 当 $0 < \alpha < 1$ 时, 正则化方法起到了加强观测而抑制背景场的作用.

2.2. 雷达观测算子

对于径向速度, 观测算子设计如下[3]:

$$v_r = u \frac{x - x_r}{r_r} + v \frac{y - y_r}{r_r} + (w - v_T) \frac{z - z_r}{r_r}, \quad (7)$$

其中, u, v, w 为速度分量, (x, y, z) 和 (x_r, y_r, z_r) 分别为雷达观测和雷达站的坐标, $r_i = \sqrt{(x - x_r)^2 + (y - y_r)^2 + (z - z_r)^2}$ 表示雷达观测和雷达站之间的径向距离, v_T 表示雨滴下落末速度[3],

$$v_T = 5.40a \cdot q_r^{0.125}, a = (p_0/\bar{p})^{0.4}, \quad (8)$$

其中, q_r 为雨水混合比, a 表示修正因子, \bar{p}, p_0 分别表示基本态气压和地面气压.

为了避免引入新的控制变量, 采用 Richardson 方程[1]提取雷达径向速度包含的大气垂直运动信息. 在连续方程的基础上引入绝热热力学方程和流体静力学方程, 得到 Richardson 方程如下:

$$\begin{aligned} \gamma \bar{p} \frac{\partial w'}{\partial z} = & -\gamma p' \frac{\partial w}{\partial z} - \gamma \bar{p} \nabla \cdot \mathbf{v}'_h - \gamma p' \nabla \cdot \bar{\mathbf{v}}_h - \bar{\mathbf{v}}_h \nabla p' \\ & - \mathbf{v}'_h \nabla \bar{p} + g \int_z^\infty \nabla \cdot (\rho \mathbf{v}'_h) dz \\ & + g \int_z^\infty \nabla \cdot (\rho' \bar{\mathbf{v}}_h) dz, \end{aligned} \quad (9)$$

其中, $\bar{\mathbf{v}}_h$ 和 \mathbf{v}'_h 分别表示水平风矢量的基本态和扰动态形式, γ 为空气定压比热和定容比热之比, g, z 和 ρ 为分别表示大气重力加速度、高度和空气密度.

假设降雨粒子满足 Marshal-Palmer 粒子分布函数($n_0 = 8 \times 10^6 \text{ mm}^{-4}$), 采用 Rayleigh 后向散射模型, 可以得到反射率因子的观测算子为[3] (单位为 dBZ),

$$Z = 43.1 + 17.5 \log(\rho q_r). \quad (10)$$

由于同化过程控制变量为总的水汽含量, 须引入暖雨参数化方案, 其主要包括水汽凝结成云水 (P_{COND})、雨水碰并云水 (P_{RACW})、云水向雨水的自动转换 (P_{RAUT}) 和雨水的蒸发 (P_{REVP}) 过程. 详细的参数化过程请参阅文献[3].

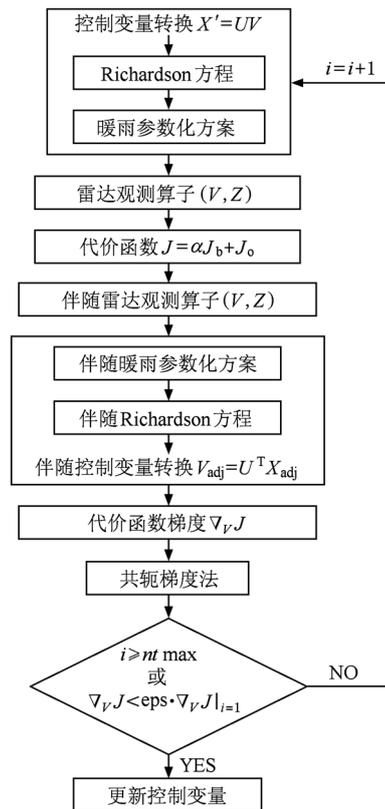


图1 正则化方法直接同化多普勒雷达资料示意图. 其中, $X' = (u', v', T', q', p'_s)$ 为模式分析变量增量, (u, v 为水平速度分量, T 为温度, q 为水汽混合比, p_s 为地表气压); $V = (\psi, \chi, T_u, p_{\text{stc}}, q/q_{\text{sg}})$ 为控制变量(干控制变量: 流函数 ψ 、势函数 χ 、非平衡温度 T_u 和地表气压 p_{stc} , q/q_{sg} 为湿控制变量); U 和 U^T 表示控制变量与分析增量之间的转化算子及其伴随形式, 且 $B = UU^T$

3. 数值试验及方案设计

3.1. 实例选取

本文针对 2008 年 8 月 14 日北京地区一次局地暴雨过程进行研究. 此次降雨从 2008081411 时(北

京时,下同)开始,于2008081420时结束.整个降雨过程呈现持续时间长、降雨强度大、地区差异大明显.个别地区出现特大暴雨,如大兴站、昌平站和渤海镇站2008081414时3h累积降雨量超过了50mm(见图2(a));李遂站和龙湾屯站2008081417时3h累积降雨量也超过了50mm(见图2(b)).位于北京

南郊的多普勒天气雷达(型号:CINRAD/SA, 116.47°E, 39.80°N)对这次强对流天气过程进行了全程观测.雷达观测采用14个仰角的VCP11体扫模式,仰角范围为0.5°—19.5°.雷达径向速度分辨率250m,有效探测范围230km;反射率因子分辨率1000m,有效探测范围460km.

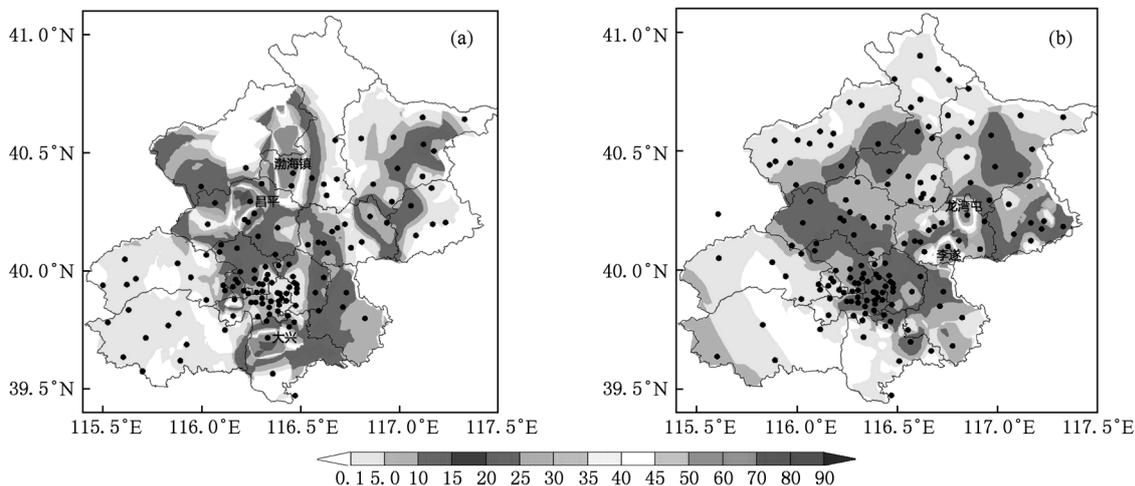


图2 3h累积实况降雨(单位:mm) (a)2008081414时;(b)2008081417时

3.2. 方案设计

本文采用的模式为WRFV3.1.1版本,而同化系统则是基于WRF模式的变分同化模块WRFDA进行设计.为了检验正则化方法的正确性,我们设计了两类试验:控制实验和同化试验.其中,同化试验于2008081408时同化一次雷达资料后进行12h预报,而控制试验从2008081408时开始冷启动,不作同化,直接进行12h预报.同时,雷达数据在同化之前,先采用VDRAS系统进行数据预处理,包括质量控制和坐标变换两部分:质量控制主要是剔除雷达数据中包含的静止地物杂波和消除速度模糊现象,同时移除雷达反射率数据存在的零度层亮带;采用Barnes插值把球坐标系下的雷达观测转换到包含经纬度信息的笛卡尔坐标系下,并对数据进行填补和平滑处理,最终将数据保存为同化所需的数据格式.

数值试验采用三重网格嵌套方案,水平分辨率为27km,9km和3km,水平格点数为151×151,184×142和192×172,垂直方向为不等距的38层(具体模拟区域请见图3);模式顶层气压为50hPa,时间积分步长为180s,60s和20s;模式物理过程选取如下:WSM6云微物理方案,YSU行星边界层方

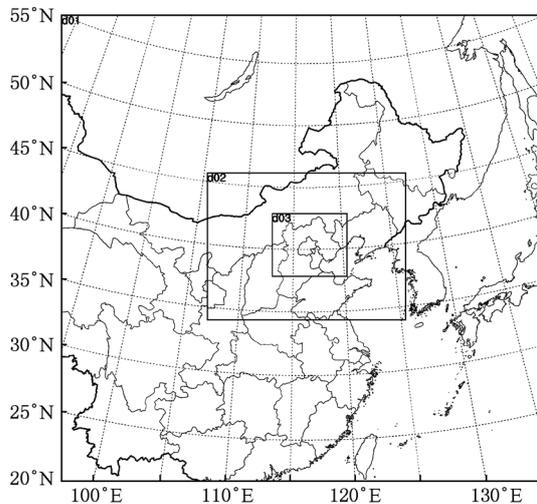


图3 模式三层嵌套区域示意图

案,RRTM长波和Dudhia短波辐射方案,其中区域一和二采用Kain-Fritsch对流参数化方案,区域三关闭对流参数化方案.模式背景场和区域一的侧边界条件均由NCEP的1°×1°再分析资料提供.同化试验的背景场误差是用NMC方法对一个月的WRF预报结果进行统计得到的气候意义上的协方差矩阵,而观测误差则简单的取为常数,径向速度1m/s,反射率因子1dBZ.

3.3. 正则化参数最优选取

正则化参数的选择是正则化方法成功的关键. 在表征观测资料 Y^o 中误差水平参数 δ 能够获取的情况下, Morozov 偏差原理 (discrepancy principle) [24] 是十分有效的一种正则化参数选择方法. 然而对于多普勒雷达观测, 由于其受地形、天气和仪器的影响非常大, 难以确定. 在变分同化系统中, 通常的做法是假定为常数, 存在很大的误差. 当误差水平参数 δ 未知时, Tikhonov 提出用拟最优准则 [11] 来确定正则化参数, 其基本思想是: 让正则化参数 α 以及对应的正则化解的变化率同时稳定在尽可能小的水平上. 在此基础上, 人们相继提出了广义交叉检验准则 [25] (generalized cross-validation, GCV) 和 L 曲线准则 [15] (L-curve). 广义交叉检验准则需要极小化 GCV 函数, 计算量非常大, 同时 α 对应的极小值

点难于求得. L 曲线准则是以 log-log 尺度来描述正则化解的模 $\|X_\alpha^\delta\|$ 及其与原始观测差值的模 $\|Y^o - H(X)\|$ 的曲线对比, 进而根据该对比结果来确定正则化参数.

为了避免观测误差 δ 不确定的影响和减少计算代价, 本文采用 L 曲线准则选取最优正则化参数. 基于同一量纲进行对比, 本文将正则化解的模 $\|X_\alpha^\delta\|$ 用 $\|X_\alpha^\delta - X_b\|$ 来代替; 所以 L 曲线即为一组正则化参数对应的 $\|X_\alpha^\delta - X_b\|$ 和 $\|Y^o - H(X_\alpha^\delta)\|$ 在 log-log 尺度下构成的图像. 根据 Engl 准则 [26]: 从逼近的角度看, 应使 α 越小越好; 然而从数值计算稳定的角度考虑, 则参数 α 取得越大越好. 我们选取了一组正则化参数 α 如下: 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.01, 0.005, 0.001; 其中 $\alpha = 1.0$ 时, 即对应传统的 3DVAR 方案. 绘成的 L 曲线见图 4(a).

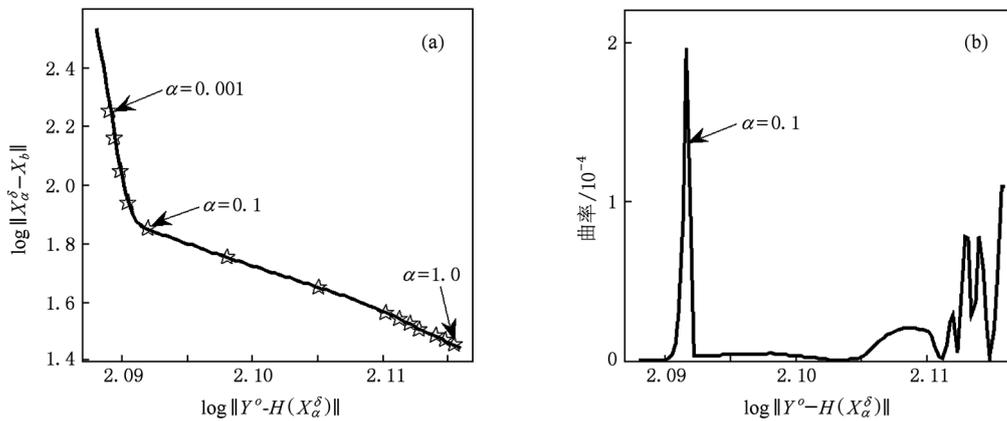


图 4 (a) L 曲线图; (b) 曲率分布图

根据 L 曲线准则, 确定最优正则化参数的关键是找到曲线的“角点”, 此时的正则化参数即为最优. Hanke [15] 定义“角点”为 L 曲线最大曲率值点. 图 4(b) 为对 14 个样本数据进行曲线拟合后得到的曲率分布. 可以看出, 曲线曲率最大值出现在 $\alpha = 0.1$ 附近. 基于以上结果, 同时对比分析正则化方法和 3DVAR 方案同化多普勒雷达资料的能力, 以及不同正则化参数对同化的影响, 我们选取控制实验 (CTL) 和三个同化试验 (VAR, R01, R001) 进行分析场和预报场分析; 其中试验 VAR 表示采用 3DVAR 方案同化多普勒雷达资料, 此时对应与正则化参数 $\alpha = 1.0$; 试验 R01 和 R001 分别表示采用正则化参数 α 取 0.1 和 0.01 的情形.

4. 同化效果研究

4.1. 误差分析

定义均方根误差 (RMSE) 如下:

$$RMSE = \sqrt{\frac{\sum_n [Y^o - H(X)]^2}{n}}, \quad (11)$$

若 X 为背景场 X^b , 则表示雷达资料同化之前, 观测和背景场之间 ($O - B$) 的均方根误差; 若 X 为分析场 X^a , 则表示雷达资料同化后, 观测场和分析场之间 ($O - A$) 的均方根误差.

表1 多普勒雷达资料同化前($O-B$)和同化后($O-A$)的均方根误差

试验	RMSE(V_r)		RMSE(Z)	
	$O-B$	$O-A$	$O-B$	$O-A$
VAR	2.7900	1.4574	2.8165	2.6247
R01	2.7900	1.2866	2.8165	2.6037
R001	2.7900	1.2436	2.8165	2.5943

表1给出了多普勒雷达径向速度 V_r 和反射率因子 Z 在同化前后的均方根误差值. 对比发现,在同化雷达资料之后,分析场和观测场之间的均方根误差均小于背景场和观测场之间的均方根误差;分析场比背景场更接近观测,达到了同化的目的. 同时,径向速度 V_r 同化前后的 $RMSE$ 变化较大,

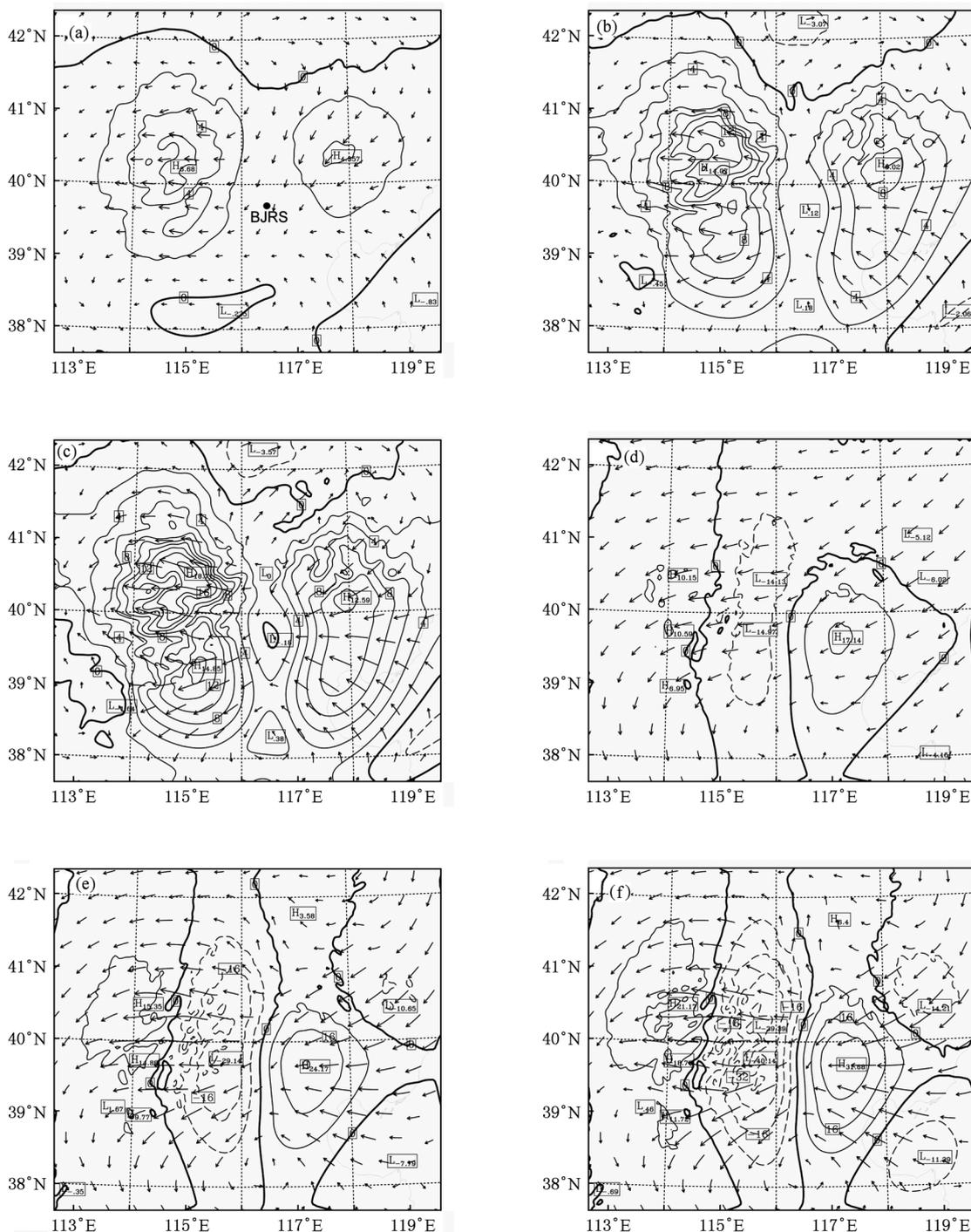


图5 850 hPa 等压面上试验(a) VAR, (b) R01 和(c) R001 的同化分析场和控制实验初始场的水平风速(间隔 2 m/s 的等值线)和风矢量(曲线箭头);850 hPa 等压面上试验(d) VAR, (e) R01 和(f) R001 的分析风场(曲线箭头),以及 700 hPa 等压面上同化分析场和控制实验初始场的垂直速度差值场(间隔 8 cm/s 的等值线);其中虚线段表示负值

同化后的值只有同化前的一半,表明径向速度 V_r 的同化方案选取得比较合理;而且正则化方法能够进一步改进同化效果, α 越小,同化后的分析场与观测场越接近. 应该指出的是,三个实验对反射率因子 Z 的同化均不理想. 分析可能的原因有:一是在反射率因子 Z 数据同化过程中,初始场中雨水混合比信息为零,简单的暖雨方案存在一个初始非平衡过程;二是同化选取的简单暖雨参数化过程,不包含冰相过程,使得同化的雷达观测限制在 5 km 以下;三是背景场误差和观测误差选取不够合理,文中背景场误差采用的是气候平均的协方差矩阵,不包含随天气过程变化的流依赖(flow-depend)信息,同时观测误差假定为常数,不符合雷达数据各向异性的特点;四是观测数目过少(只有 2667 个反射率因子观测值),观测场容易被模式噪声湮没掉.

4.2. 风场分析

试验 VAR, R01 和 R001 均在 2008081408 时同

化一次雷达资料,其初始猜值场即为控制实验 CTL 的初始场. 图 5 给出了三个同化试验同化雷达资料后得到的分析场相对于控制实验初始场在 850 hPa 等压面上的水平风速和风矢量的差值分布情况,以及在 700 hPa 等压面上垂直速度的差值场. 同化雷达资料后,三个实验的水平风场均有明显的变化(见图 5(a),(b)和(c)),而且随着 α 减小,水平风场变化的幅度和范围相应变大,其中实验 R001 的最大风速增量达到了 18 m/s,影响半径达到了 400 km. 同时试验 R01 和 R001 的分析场还存在负增量区,去掉了一些虚假强对流运动,而且试验 R01 和 R001 的气旋性环流增量比试验 VAR 要明显,增强了辐合辐散运动,有利于暴雨的发生和发展. 图 5(d),(e)和(f)的垂直速度差值场表明同化实验成功提取了雷达观测中包含的垂直速度信息. 垂直速度的变化主要是在雷达东边加强了上升运动,而在雷达西边加强下沉运动,三个实验最大的垂直速度改变量分别为 17.14 cm/s (VAR), 24.17 cm/s

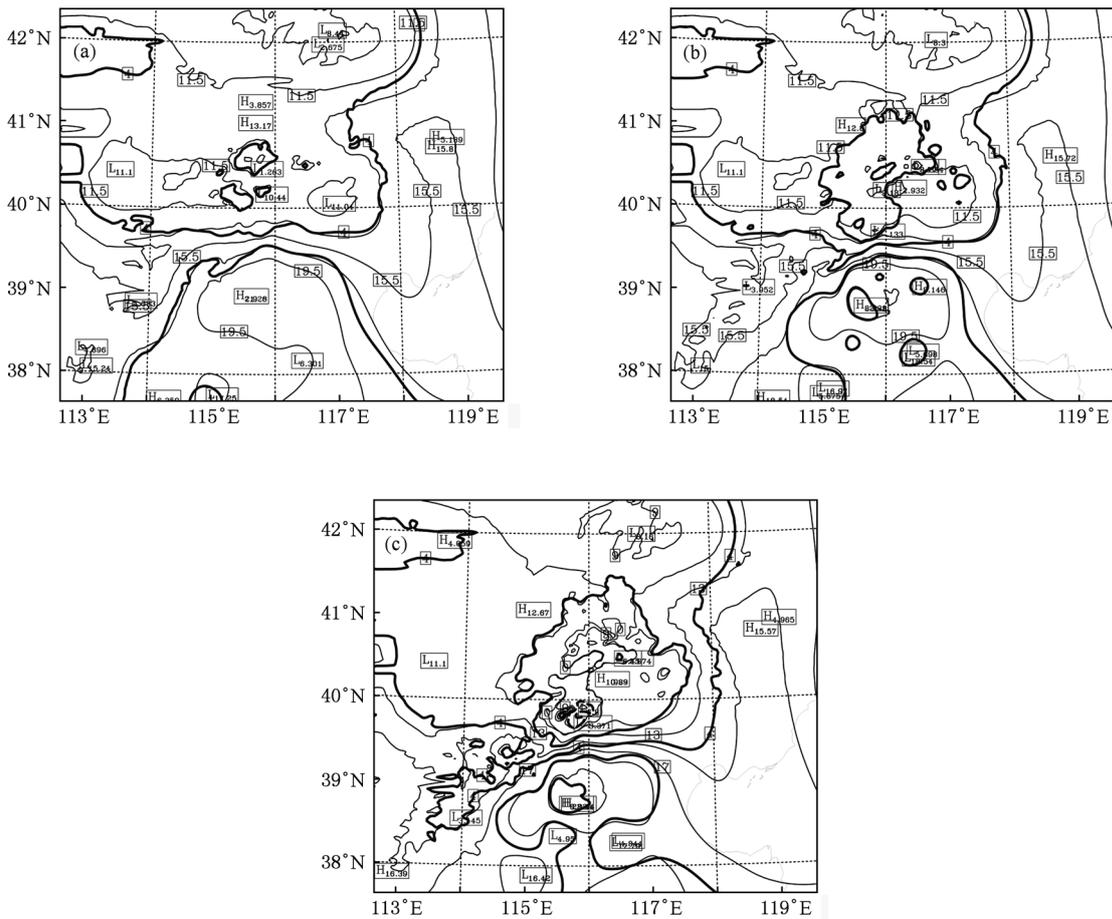


图 6 850 hPa 等压面上试验(a) VAR, (b) R01 和(c) R001 同化后的水汽场(细等值线,间隔 2 g/kg),以及同化分析场和控制实验初始场的水汽差值场(粗等值线,间隔 2 g/kg);其中虚线段表示负值

(R01) 和 31.68 cm/s (R001). 由以上的分析可知, 选取不同的正则化参数会导致不同的风场同化结果, 而且随着正则化参数 α 的减小, 同化引入的中小尺度风场观测信息就越多, 风场结构越细致.

4.3. 质量场分析

Xiao 等^[1,2] 指出同化雷达径向速度改变显著的是动力场, 而同化反射率数据主要调整的是质量场 (热力场和水汽场). 图 6 给出了 850 hPa 等压面上三个同化试验同化得到的水汽分布及其相对于控制实验初始场的差值场. 分析可知, 三个分析场的水汽差值幅度随正则化参数 α 变化不大, 最大水汽增量依次为 6.301 g/kg (VAR), 8.238 g/kg (R01) 和 8.935 g/kg (R001); 但是在发生了强降雨的区域, 如大兴站、昌平站等, 水汽增量场的结构有较大不同, 其中以实验 R01 的水汽增量结构与实况降水分布最为相似. 从水汽场来看, 三个实验的分析场大致相同, 这也与上述的误差分析结果符合. 同时三个同化实验得到的热力场和控制实验的差值分布与水汽差值场结构相似 (图略).

总的来说, 不论是风场还是质量场, 其同化效果均与正则化参数 α 的选取有关. 一般 α 越小, 同化的中小尺度观测信息就越多, 得到的分析场与观测越接近. 同时, 同化效果的好坏与背景场和观测的质量、背景误差和观测误差的好坏、数值模式等有关, 并不能单单由正则化参数 α 来决定.

5. 降雨预报评估

5.1. 降雨模拟

图 7 给出了数值试验在 2008081414 时北京地区 3 h 累积预报降雨分布. 由图 7(a) 可知, 控制试验 CTL 正确预报出了大兴站附近的强降雨区, 预报强度相对实况要弱, 同时预报对朝阳区奥林匹克公园附近的大雨也有所描述, 但是实验对其他地区的强降雨中心和雨带基本上没有做出预报. 图 7(b) 表明, 在采用 3DVAR 方案同化雷达资料之后, 实验 VAR 在 2008081414 时的 3 h 累积降雨预报要好于控制实验 CTL, 试验基本上正确预报出了城南大兴站的强降雨中心的位置和雨强, 同时对怀柔渤海镇到九渡河站的强降雨带以及顺义龙屯湾站的降雨中心也有刻画, 但是实验还是没有能够预报出昌平

地区的强降雨, 同时城区的降雨分布与实况还是较大出入. 从图 7(c) 可知, 试验 R01 不仅正确预报出了大兴站地区的强降雨中心和渤海柔地区强降雨带, 还正确刻画了密云北庄地区和丰台鹰山公园附近的强降雨, 此时昌平地区的强降雨中心终于被预报出来, 强度与实况也比较接近, 同时实验 R01 还是漏报了龙屯湾地区的强降雨. 图 7(d) 表明, 试验 R001 正确预报出了 2008081414 时昌平站和大兴站两个强降雨中心, 但是强度相对实况偏弱, 同时龙屯湾站和九渡河站附近降雨也有所体现, 整体的雨区分布与实况比较接近.

为了分析雷达资料同化后的预报时效, 我们给出了各试验在 2008081417 时的 3 h 累积预报降水分布图 (见图 8). 从图 8(a) 来看, 此时控制实验 CTL 很好地预报出了顺义地区龙湾屯、杨镇到李遂站的强降雨带和密云地区的降雨, 但是强度相对实况要强; 同时实验对城区的降雨也有所刻画, 但是整个雨带的分布与实况存在较大的差异, 而且平谷黑豆峪和夏各庄的强降雨区实验也没有预报出来. 同时图 8(b) 的结果表明实验 VAR 在 2008081417 时的预报并不够理想, 与实况相差较大. 可以看出, 除了顺义龙湾屯到杨镇站的强降雨中心和密云县中部的雨区以及城区奥林匹克公园到三间房站的雨带被预报出来以外, 平谷的强降雨区以及大部分的降雨区域均没有预报出来. 令人欣慰的是, 实验 R01 在此时预报效果依然比较好 (见图 8(c)), 除了夏各庄和黑豆峪站的强降雨区没有预报出来以外, 龙湾屯、杨镇到李遂站的强降雨带, 城区十八里店、三间房和奥林匹克公园的降雨区都被正确预报出来, 但是整体的降雨分布与实况还有一定的偏差. 而图 8(d) 的结果则表明此时试验 R001 的预报不如 R01, 除了龙湾屯、杨镇到李遂强降雨带和黑豆峪到夏各庄的降雨区与实况相符外, 通州和丰台的两个强降雨中心均为虚假降水, 而且其强度远远超过了实际降雨量.

从以上的分析结果可知, 控制实验 CTL 对局地强暴雨的预报能力还有待提高, 特别是在模式起转初期, 由于用大尺度模式预报结果 (NCEP 再分析资料) 生成的初始场, 无法提供中小尺度天气系统信息, 模式存在起转延迟现象 (spin-up), 从而导致降雨预报存在一定的延迟; 在经过一段时间运行调整后, 模式达到平衡, 预报效果开始改善, 但是降雨预报强度大于实况. 而采用 3DVAR 方案同化多普勒天气雷达资料后, 模式初始场中加入了中小尺度天

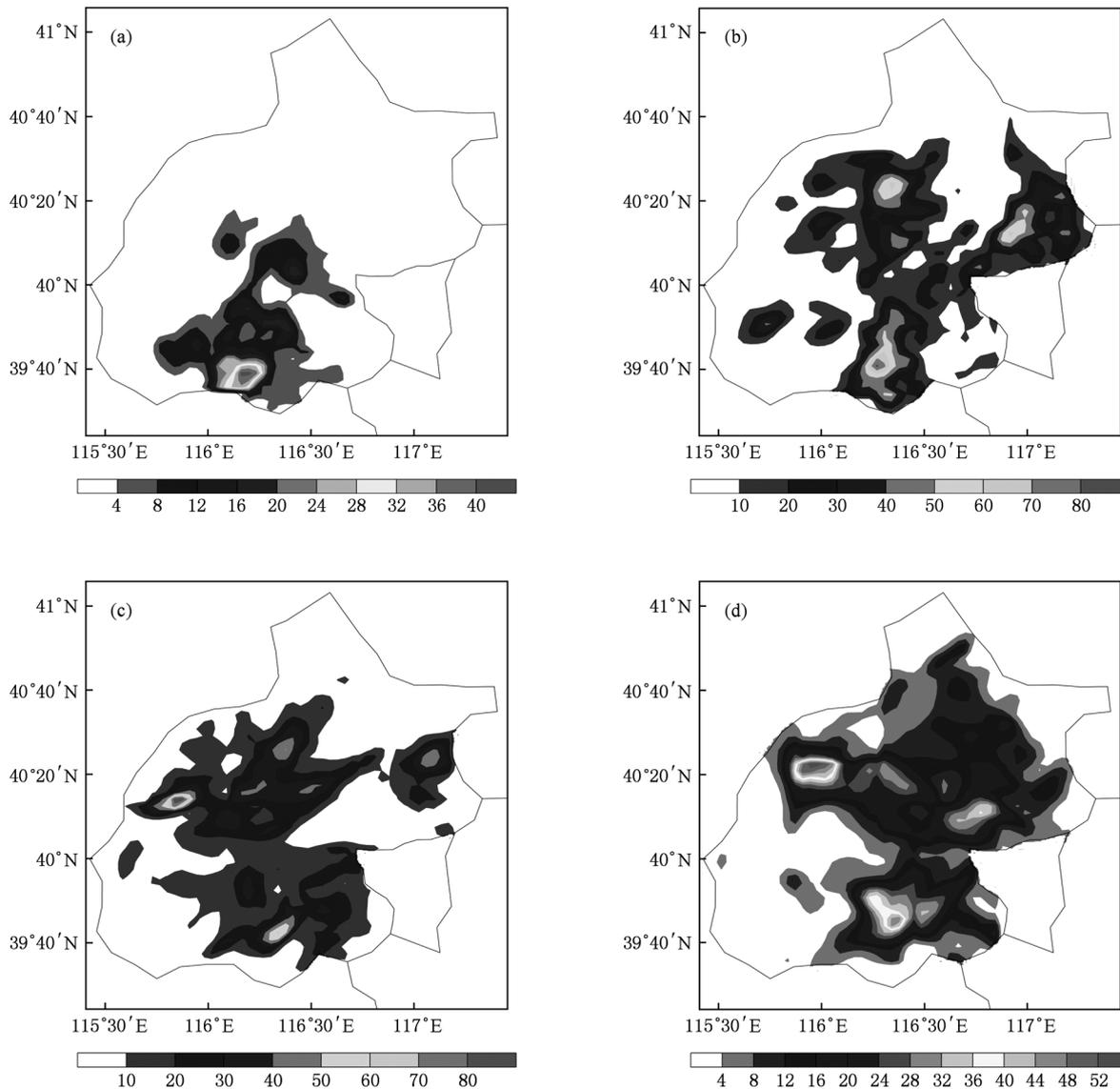


图7 2008081414 时的 3 h 累积预报降水预报 (a) 试验 CTL; (b) 试验 VAR; (c) 试验 R01; (d) 试验 R001

气结构信息,能够在一定程度上缩短了模式起转调整的时间,从而提高了模式初期降雨预报的准确度.但是经过一段时间的积分后,雷达资料的效果消失,降雨预报效果开始下降.而对试验 R01 和 VAR 的降雨预报结果分析表明, α 取 0.1 时的正则化参数更能提取雷达资料中的有效信息,合理分析场中的流场结构和水汽分布,相应延长雷达资料同化时效,从而改进了整体的降雨预报结果.同时试验 R001 的结果表明,选取不同的正则化参数会导致不同的预报效果.随着正则化参数 α 的减小,模式初始场与观测场更加接近,但是试验 R001 的降雨预报效果并没有好过试验 R01. 这可能是由于雷达观测中本身就存在误差和虚假回波信息,所以若在

分析场中过度地引入中小尺度观测信息反而会增加模式初始场的噪声,破坏模式初始场的平衡,不利于对流系统的发展,从而导致预报变差;同时数值模式中描述强对流性天气的物理过程和参数化方案仍然不够完美也会对降雨预报产生影响.

5.2. 定量降水预报

为了进一步定量检验短时降雨预报的效果,我们引入 QPF 的 TS 评分技巧,

$$TS = \frac{C}{F + O - C},$$

其中, O 为实际发生的降雨事件, F 为预报的降雨事件, C 则表示被正确预报的降雨事件. 具体的计

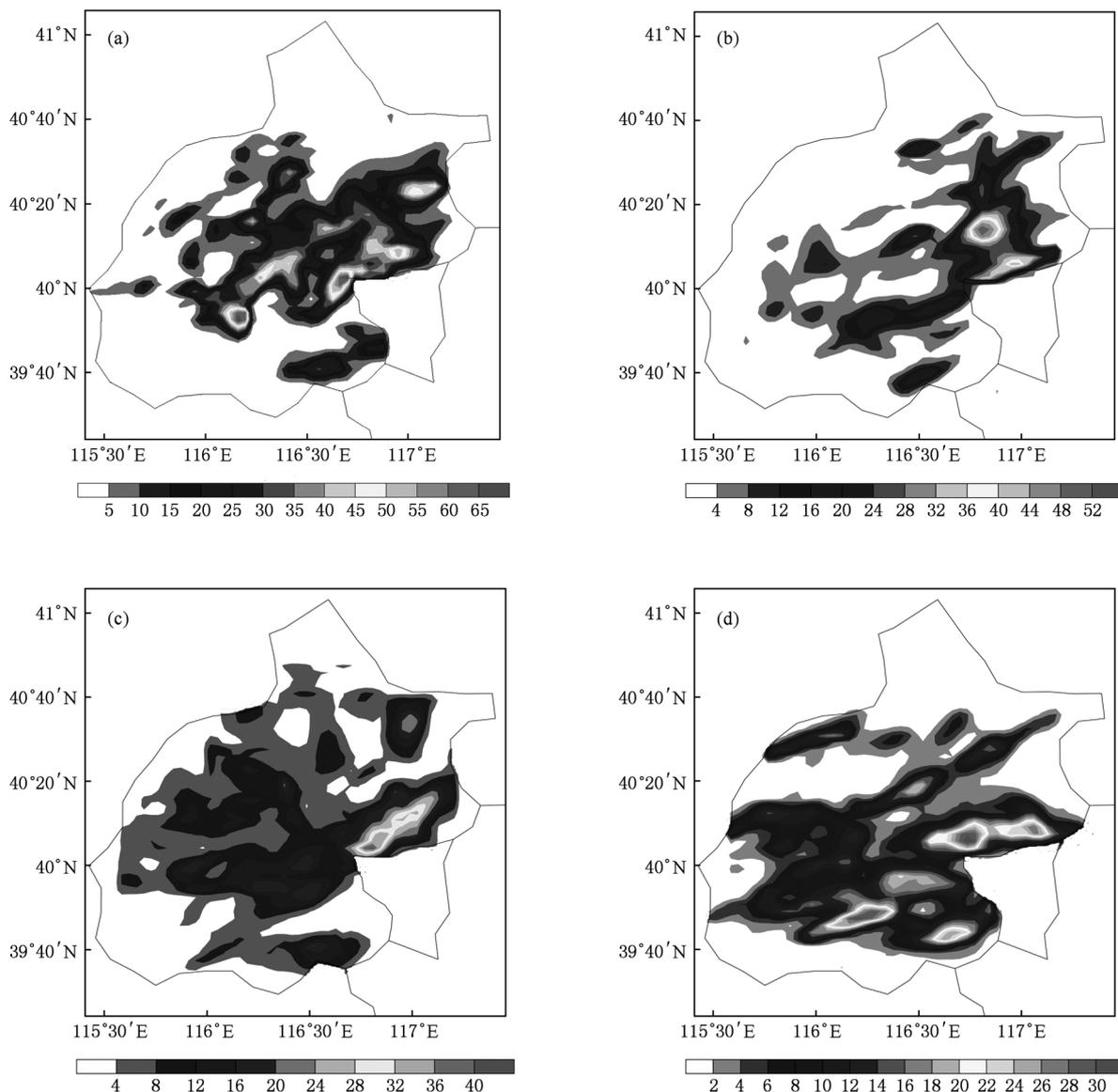


图8 20080814时的3 h 累积预报降水预报 (a) 试验 CTL; (b) 试验 VAR; (c) 试验 R01; (d) 试验 R001

算方法如下: 首先把实况降雨量当做一个集合, 给定降雨阈值 Th , 则 O 等于实况降雨大于 Th 的站点集合, F 等于预报降雨量结果大于 Th 的站点集合, C 等于实况和预报结果均大于 Th 的站点集合. 对 CTL, VAR, R01 和 R001 四个实验在 20080814 时的 1 h 和 3 h 累积预报降水作 0.1 mm, 5 mm 和 10 mm 降水的 TS 评分, 观测降水资料来自于北京地区的 217 个地面自动站资料, 得到的 TS 评分如图 9.

分析图 9 的结果可知, 在所有试验中, R01 的 TS 评分最好, R001 和 VAR 次之, 控制实验 CTL 最差, 符合上述分析结果. 同时可以看出, 在各种阈值下, 3 h 累积预报降雨 TS 评分整体比 1 h 累积预报降雨 TS 评分要好, 表明雷达资料的同化有一定的

时效, 一般在 6 h 左右. 同时我们也应指出, 当阈值大于 10 mm 时, 各试验的 TS 评分结果都较低, 这说明模式对局地暴雨的预报还是不够理想, 同时同化过程采用的简单暖雨过程与预报模式的 WSM6 云微物理方案的差异, 也会影响降雨预报结果.

6. 总结与讨论

本文首先分析了直接同化多普勒雷达资料的 3DVAR 理论. 在此基础上, 提出用数学物理反问题中的正则化方法直接同化多普勒天气雷达资料. 在未知观测误差的情况下, 采用 L 曲线方法进行最优正则化参数的选取. 为了对比正则化方法和 3DVAR

方案同化多普勒雷达资料的效果,以及选取不同的正则化参数对同化的影响,总共选取了四个数值试验:控制实验 CTL 和三个同化试验 VAR, R01 和

R001 进行对比分析. 我们从同化误差、分析场、降雨模拟,以及 TS 评分四个方面对正则化方法进行评估,主要得到以下几点结论:

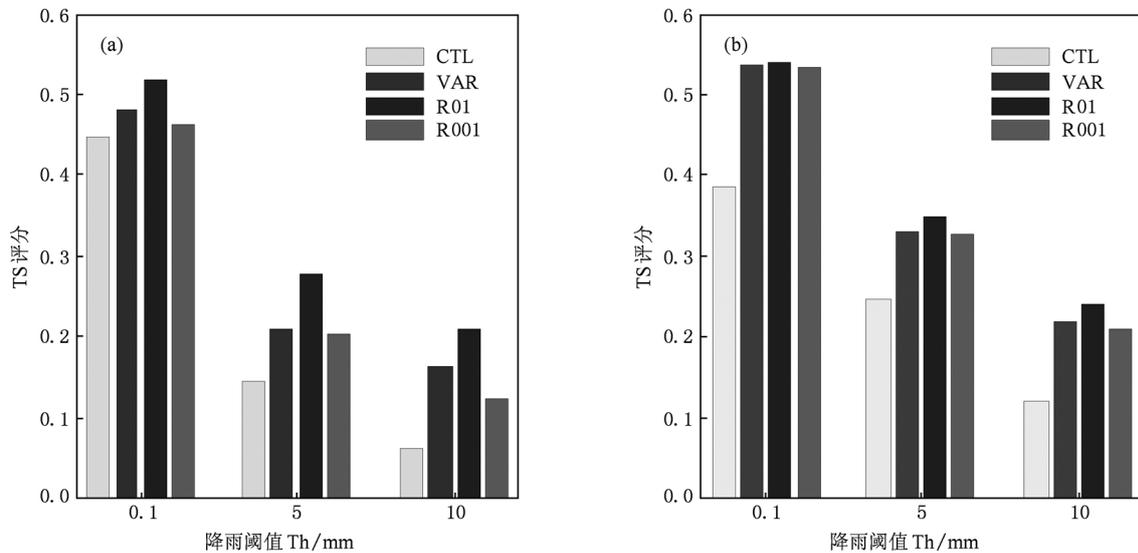


图9 2008081414 时各试验 (a)1 h 和(b)3 h 累积预报降雨的 TS 评分

1. 正则化方法和 3DVAR 方案均能有效同化多普勒雷达资料,而且雷达径向速度的同化效果明显好于反射率因子,同化后的径向速度 RMSE 相比于同化前减少了将近一半,而反射率因子 RMSE 在同化前后变化不大.

2. 正则化参数 α 越小,分析场各变量的增量越明显,同化进入初始场的中小尺度信息也越多,分析场与观测场也就越接近,其中风场的改变量相比质量场(水汽含量)要明显.

3. 采用不同正则化参数 α 同化多普勒雷达资料,均能在一定程度上缩短模式起转调整时间,提高短期降雨预报的准确率;其中最优选取正则化参数 α 的同化试验能够延长雷达资料同化时效,进一步改善降雨预报效果,合理雨区雨带分布,相应得到的 TS 评分也就更高.

4. 正则化参数 α 的选取,对同化效果至关重要. 若 α 选取过大,高分辨率的雷达观测资料就会被

背景场湮没掉,起不到应有的作用;若 α 选取过小,在模式初始场中过度地引入中小尺度观测信息反而会增加模式初始场的噪声,破坏模式初始场的平衡,导致预报变差.

从本文的分析来看,正则化方法在理论和实践上均有一定意义,适用于雷达资料同化. 但是在本文中,通过事先给定一组正则化参数,然后采用 L 曲线方法进行最优正则化参数的选取,会造成计算资源不必要的浪费,不适用于业务运行. 因此,如何有效确定雷达观测的误差水平,进而采用 Morozov 偏差原则或模型函数方法寻找到最优的正则化参数是下一步研究的重点.

感谢美国国家大气研究中心 (NCAR) 提供的 WRF 模式,感谢中国气象局北京城市气象研究所范水勇研究员提供的多普勒天气雷达资料.

[1] Xiao Q N, Kuo Y H, Sun J Z, Lee W C, Lim E, Guo Y R, Barker D M 2005 *J. Appl. Meteor.* **44** 768
 [2] Xiao Q N, Kuo Y H, Sun J Z, Lee W C, Barker D M, Lim E 2007 *J. Appl. Meteor. Climatol.* **46** 14
 [3] Sun J Z, Crook N A 1997 *J. Atmos. Sci.* **54** 1642

[4] Sun J Z, Crook N A 1998 *J. Atmos. Sci.* **55** 835
 [5] Sun J Z, Crook N A 2001 *Wea. Forecasting* **16** 117
 [6] Tong M, Xue M 2005 *Mon. Wea. Rev.* **133** 1789
 [7] Wan Q L, Xue J S, Zhuang S Y 2005 *Acta Meteorologica Sinica* **63** 129 (in Chinese) [万齐林, 薛纪善, 庄世宇 2005 气象学报

- 63 129]
- [8] Yang Y, Qiu C J, Gong J D, Huang J 2008 *Acta Meteorologica Sinica* **66** 479 (in Chinese) [杨毅、邱崇践、龚建东、黄静 2008 气象学报 **66** 479]
- [9] Sheng C Y, Pu Y F, Gao S T 2006 *Chinese Journal of Atmospheric Sciences* **30** 93 (in Chinese) [盛春岩、浦一芬、高守亭 2006 大气科学 **30** 93]
- [10] Xu X Y, Liu L P, Zheng G G 2006 *Chinese Journal of Atmospheric Sciences* **30** 712 (in Chinese) [许小永、刘黎平、郑国光 2006 大气科学 **30** 712]
- [11] Thikhonov A. N 1963 *Dokl. Akad. Nauk. SSSR* **151** 49
- [12] Huang S X, Wu R S 2005 *Mathematical and physical problem in atmospheric science* (Beijing: Meteorological Press) p422—433 (in Chinese) [黄思训、伍荣生 2005 大气科学中的数学物理问题(北京:气象出版社) 第422—433页]
- [13] Sheng Z, Huang S X 2010 *Acta Phys. Sin.* **59** 1734 (in Chinese) [盛 峥、黄思训 2010 物理学报 **59** 1734]
- [14] Huang S X, Cao X Q, Du H D, Zhong J Q 2010 *J. of Hydrodynamics* **B 22** 185
- [15] Hanke M, Hansen P C 1993 *Surv. Math. Ind.* **3** 253
- [16] Kalnay E 2003 *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge: Cambridge University Press) p168
- [17] Cardinali C, Pezzulli S, Anderson E 2004 *Q. J. R. Meteorol. Soc.* **130** 2767
- [18] Huang S X, Sheng Z 2006 *Acta Phys. Sin.* **55** 6720 (in Chinese) [黄思训、盛 峥 2006 物理学报 **55** 6720]
- [19] Cai Q F, Huang S X, Gao S T, Zhong K, Li Z Q 2008 *Acta Phys. Sin.* **57** 3912 (in Chinese) [蔡其发、黄思训、高守亭、钟 科、李自强 2008 物理学报 **57** 3912]
- [20] Zhang L, Huang S X, Liu Y D, Zhong J 2010 *Acta Phys. Sin.* **59** 2889 (in Chinese) [张 亮、黄思训、刘宇迪、钟 剑 2010 物理学报 **59** 2889]
- [21] Huang S X, Zhao X F, Sheng Z 2009 *Chin. Phys. B* **18** 5084
- [22] Kress R 1989 *Linear Integral Equation* (New York: Springer-Verlag) p250—258
- [23] Kirsch A 1996 *An Introduction to Mathematical Theory of Inverse Problem* (New York: Springer-Verlag) p48—53
- [24] Morozov V A 1966 *J. Comp. Math. Math. Phys.* **6** 170
- [25] Golub G H, Heath M, Wahba G 1979 *Technometrics* **21** 215
- [26] Engl H W, Gfrerer H 1988 *Appl. Numer. Math.* **4** 395

Regularization method of assimilating Doppler radar data and its influence on precipitation forecast*

Zhao Yan-Lai^{1)†} Huang Si-Xun^{1)‡} Du Hua-Dong¹⁾ Zhong Ji-Qin²⁾

1) (Institute of Meteorology, PLA University of Science and Technology, Nanjing 211101, China)

2) (Institute of Urban Meteorology, China Meteorological Administration, Beijing 100089, China)

(Received 21 May 2010; revised manuscript received 26 September 2010)

Abstract

A new regularization method is proposed to directly assimilate Doppler radar data into mesoscale numerical weather forecast based on the traditional 3DVAR. For seeking the minimum module solution of the $Y^n = H(X)$ with bias δ , the regularization method is adopted and leads to a new cost function. A group of experiments were designed to study the case of a locally strong rainstorm occurred in Beijing area on August 14, 2008. The L-curve principle is used to determine the optimal regularization parameter and the result is $\alpha = 0.1$. Numerical results demonstrate that both regularization method and 3DVAR scheme can efficiently assimilate the Doppler radar data, and that the improved initial condition can alleviate the spin-up phenomenon and improve the nowcasting precipitation forecast. However, when an optimal regularization parameter is chosen, better improvement, more accurate precipitation forecast and higher TS score are expected.

Keywords: Doppler radar data, regularization methods, 3DVAR, precipitation forecast

PACS: 92.60.Wc, 92.60.Qx

* Project supported by the National Key Technology Research and Development Program of the Ministry of Science and Technology of China (Grant No. 2008BAC37B03).

† E-mail: zhaoyanlai@ gmail.com

‡ Corresponding author. E-mail: huangxp@ yahoo.com.cn